Data Viz Proposal

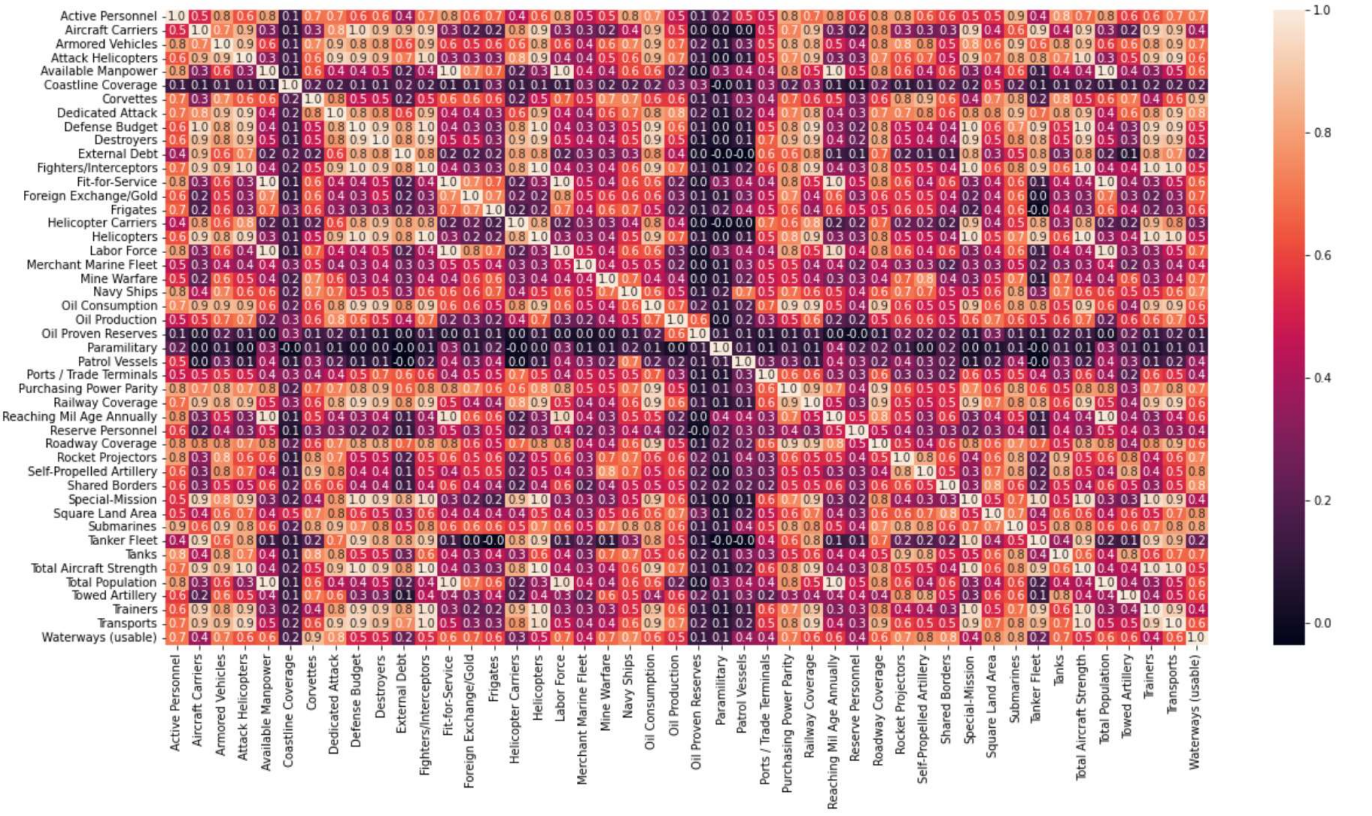
Code

By: Stuart Gavidia, Andrew Sandoval, and Kirill Perfiliev

1. Sasakitetsuya, T. (2022, February 23). Top 20 defense budget countries analysis. Kaggle.

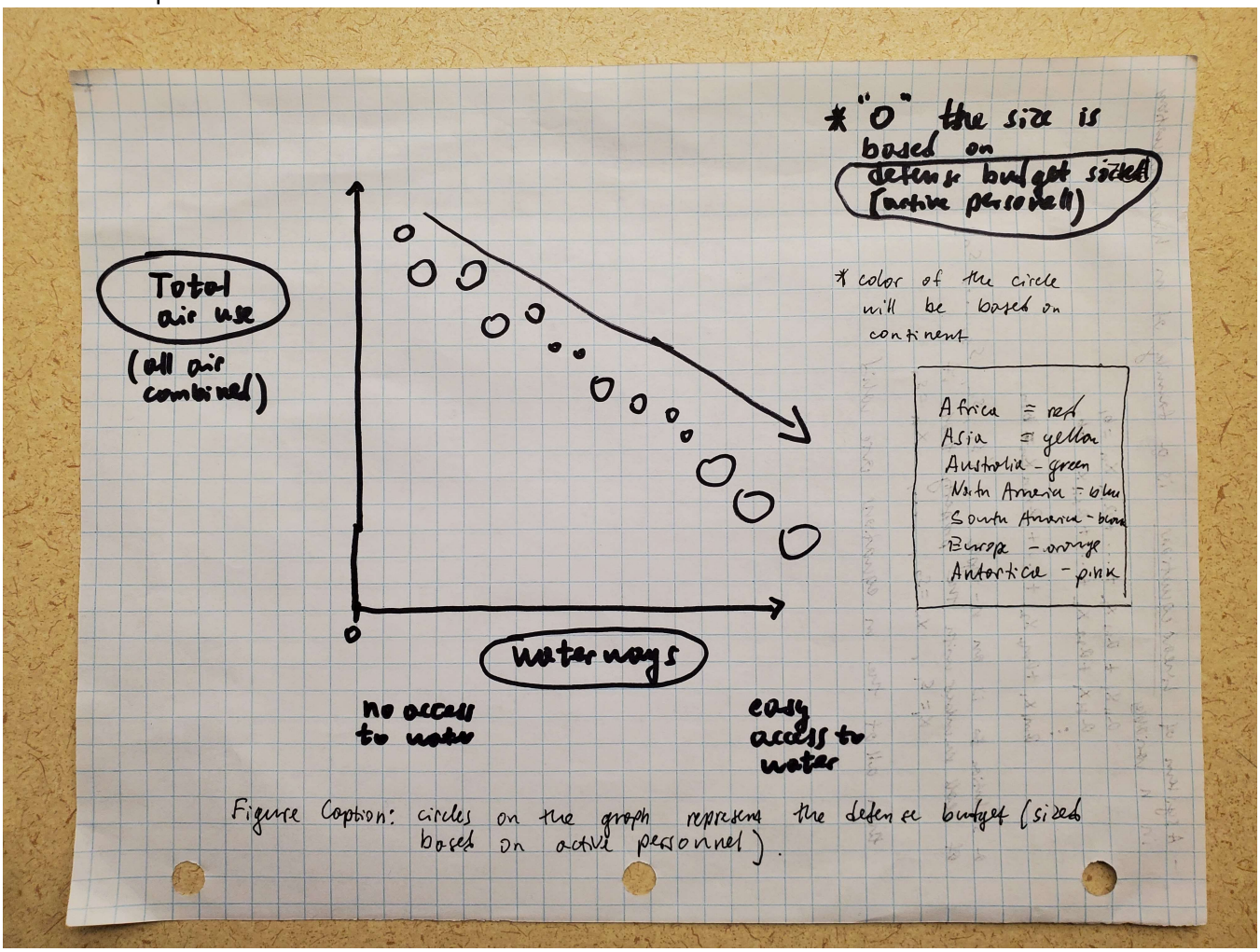
a. https://www.kaggle.com/code/sasakitetsuya/top-20-defense-budget-countries-analysis (https://www.kaggle.com/code/sasakitetsuya/top-20-defense-budget-countries-analysis)

2. Initial Graph



3. The graph in question is a heatmap, which is a data visualization tool used to represent the correlation coefficients between various features in a matrix form. In the context of military defense analysis, the features represent different aspects of military capability, such as personnel numbers, equipment inventories, and budget allocations. The colors in the heatmap indicate the strength and direction of the correlation between pairs of features: red hues suggest a positive correlation, blue hues indicate a negative correlation, and the intensity of the color corresponds to the magnitude of the correlation. For example, a deep red cell would indicate a strong positive relationship between the two variables it intersects, meaning they tend to increase or decrease together. In contrast, a deep blue cell would suggest a strong negative relationship, indicating that as one variable increases, the other tends to decrease. The purpose of this graph is to identify and visualize the degree to which different aspects of military strength are interrelated, which can inform strategic planning and resource allocation decisions.

4. Sletched Graph



5. What's the reliance on air (air power) based on the relationship between distance from the waterways and defense budget (calculated based on the active personnel) for each country?

a. We are trying to see that the further the country is away from water, the more reliant it is on air. Rather if the country has easy access to water, then there is a split in budget for air, ground, and sea

6. We are following Gelman's advice. For example, we are comparing the use of air vertically, not horizontally. We use a common x-axis. Potentially, what we haven't thought about is isolating individual comparisons of interest (creating additional graphs while properly adjusting the y-axis).

a. We are working with four variables:Countries combined into continents, Total Air Force (attack helicopters, helicopters, total aircraft strength = fighters/intercepters), Defense budget (based on the active personnel), Waterways

7. We'll include a caption that explains what each bubble in the graph represents (country) and what the size means (defense budget)

8. We'll be using the sizing of each bubble to represent the amount of defense budget that each country uses, this follows the area principle that Sharpe explained where the size of each data point corresponds to the magnitude that data represents.